DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

CREDIBILITY AND LAW ENFORCEMENT

Jennifer F. Reinganum and Louis L. Wilde

SOCIAL SCIENCE WORKING PAPER 604

April 1986

ABSTRACT

The precommitment approach to law enforcement is inappropriate as a positive theory of crime and punishment because it is inconsistent with the institutional structure of U.S. law enforcement. We develop a formal model which integrates theories of optimal sanctions, individual criminal behavior and the allocation of effort to apprehension, and imposes credibility constraints on the choice of sanction--i.e., given the severity of a crime and the individual characteristics of the criminal, the sanction imposed must be optimal from society's perspective, after the crime has been committed.

CREDIBILITY AND LAW ENFORCEMENT

Jennifer F. Reinganum and Louis L. Wilde*

> We give them pleas because we know the judge is going to give them a light sentence anyhow. The judge does that because he knows they're going to be out anyhow, from the department of corrections. I just don't think anyone really has the ultimate responsibility and looks at the social consequences of what's going on in the criminal justice system. (Barry Zavah, quoted in Jackson, 1984, p. 295)

1. INTRODUCTION

Gary Becker's seminal work on the economics of crime and punishment (1968) has had an enormous and lasting impact on the way in which we view public policy toward criminal behavior. Any analysis of criminal behavior and policy responses to it, whether in the disciplines of sociology, psychology, philosophy or law, is compelled to discuss Becker's formidable arguments (if only to describe the author's reasons for rejecting his approach; e.g., Wittman, 1974; Stern, 1978). Researchers in law and economics have adopted the approach wholeheartedly and have generalized it in several directions (Harris, 1970; Brown and Reynolds, 1973; Block and Heineke, 1975; Heineke, 1975; Polinsky and Shavell, 1979). Although Becker's theory was essentially normative, the additional hypothesis that society is indeed acting in an optimal manner generates a positive theory of criminal behavior and law enforcement. An excellent survey of this work and empirical applications of it can be found in Pyle (1983).

While Becker's approach has much to recommend it, particularly the fundamental idea that crime is amenable to economic analysis, it has one important drawback for the purposes of positive analysis: it supposes that society chooses its enforcement policy to maximize a measure of social welfare, anticipating potential criminals' optimal behavior in the face of that policy. That is, society precommits itself to an enforcement policy. The relationship between the "supply of offenses" and a given enforcement policy has come to be known in this literature as the "deterrence function." The reason why such an approach may be inappropriate as a positive theory of crime and punishment, at least for the case of the United States, is that society determines guilt or innocence, as well as the extent of sanctions, after a crime has been committed.[1] Thus an enforcement policy which appears optimal before the commission of a crime will generally not be optimal after the crime has been committed. In the language of dynamic optimization, such a policy will not be "dynamically consistent" (Kydland and Prescott, 1977); alternatively, in the language of game theory, such a policy will not be "subgame perfect" or "credible" (Selten, 1975; Kreps and Wilson, 1982a).

We have elsewhere applied the notion of sequential rationality

to a game-theoretic analysis of tax compliance (Graetz, Reinganum and Wilde, 1986), and it is our purpose in this paper to begin exploring its implications in other legal contexts. The role of credibility constraints for law enforcement is illustrated most easily in the case of an offense which is punished by imprisonment. Assuming that all individuals are potentially deterrable and that crime is socially inefficient,[2] the ideal precommitment policy is to deter all crime by employing an arbitrarily small probability of apprehension and an arbitrarily long sentence. However, prison sentences impose a variety of costs upon society. Thus it is reasonable for criminals to doubt whether society would actually carry out the threat of a very long sentence given they are unlucky enough to be apprehended. In other words, there is a credibility problem. Thus the existing institutional structure, which allows for substantial ex post flexibility in the determination of sanctions, will reduce the extent of deterrence which is available to society. In order to develop a positive theory of crime and punishment in light of this institutional structure, it is necessary to incorporate credibility constraints.

In this paper we develop a formal model which integrates theories of optimal sanctions, individual criminal behavior and the allocation of effort to apprehension, and imposes the sort of credibility constraints discussed above. Specifically, this will require that given the severity of a crime and the individual characteristics of the criminal, the sanction imposed must be optimal from society's perspective, after the crime has been committed. Under

plausible assumptions about social utility, we find that the subgame perfect equilibrium sentence schedule varies directly with the severity of the crime and inversely with a measure of the opportunity cost of incarceration. Unlike the ideal precommitment policy, sentence length is independent of the probability of apprehension. Anticipation of the equilibrium sentence schedule can result in a nonmonotonic relationship between the criminal's opportunity cost of incarceration and the equilibrium severity of his preferred crime.

If all individuals are potentially deterrable (and crime is socially inefficient), the precommitment approach implies no crime in equilibrium, while our model incorporating ex post discretion in the setting of sanction levels implies crimes of nontrivial severity. Thus ex post discretion limits the extent of deterrence. However, the requirement that the sanction policy be credible does not imply that society has no instrument of deterrence at all. The equilibrium sanction policy deters in the sense that it reduces the severity of crime. Moreover, the probability of apprehension is an important determinant of the severity of crime; increasing this probability also induces individuals to commit less serious crimes.

There are both formal and informal ways in which some types of precommitment can improve society's well-being, when used appropriately. First, the legislation of mandatory penalties can improve welfare; however, naive efforts to enhance deterrence (such as the often suggested "solution" of a small probability of apprehension accompanied by an extremely large penalty) will be at least partially

thwarted by society's refusal to convict. Some evidence for this phenomenon, known as "jury nullification," can be found in Kalven and Zeisel (1966, Ch. 21), Wilson (1975, p. 187) and Jackson (1984, p. 152-53). Second, the appointment or election of judges with particular preferences is another means of precommitment. Third, society may find it possible to precommit through informal means which rely on the fact that law enforcement is an ongoing process. For instance, the use of "trigger strategies" or the desire to establish a "reputation for deterrence" might sustain an outcome which would not be credible absent repetition of the game. We argue that trigger strategies and reputation-building are unlikely to be both feasible and optimal, at least for sustaining the ideal precommitment policy. However, the higher the probability of apprehension, the lower is the fully deterring sentence. Thus maintaining a higher probability of apprehension might also enhance the sustainability of the fully deterring policy via the use of trigger strategies or reputation-building.

In Section 2, we describe the extent of discretion enjoyed by judges and juries. In Section 3 we present our basic model and findings, and Section 4 contains an illustrative example. Section 5 considers the allocation of resources to apprehension. Section 6 evaluates the feasibility and optimality of various deterrence--enhancing strategies, including the use of mandatory sentences and the possibility of rendering credible the deterring policy through repeated play. Section 7 summarizes our results, discusses related

literature and suggests avenues for future research, including empirical implications of our approach.

2. THE EXTENT OF DISCRETION

Our analysis is based upon the observation that as representatives of society police, prosecutors, judges and juries enjoy considerable discretion in the extent to which they enforce the law. In this analysis, we will be specifically concerned with the discretionary power of judges and juries to determine conviction and the extent of sanctions. Whether sanctions are determined by judges or juries depends on the jurisdiction and the crime. Some crimes automatically receive bench trials. When juries are used, in some jurisdictions (mostly in the South) the jury determines both guilt or innocence and the extent of sanctions. In most states and in Federal cases, the jury convicts or acquits and the judge determines sanctions. However, juries often have some control over sentencing by their ability to recommend leniency or to reduce the charge, particularly in cases which involve questions of intent or the use of the death penalty. For example, Kalven and Zeisel (1966, p. 59-60) remark that "In a fair number of cases more than one charge is presented to the jury. . . . The multiple charge may either be concurrent, such as assault and carrying a concealed weapon, where the defendant may be convicted on either or both of the charges; or the multiple charge may arise from a doubt as to whether the defendant committed only a 'lesser included offense,' such as manslaughter

instead of murder. Here the defendant can be found guilty of only one of these crimes, not of both. Lesser offenses are included most frequently where a specific intent is at issue, e.g., intention to kill as against intention only to harm."

Although criminal statutes often specify allowable sanctions, they typically leave a great deal of discretion to whomever ultimately imposes the sanction. For example, in Illinois the Unified Correction Code classifies felonies into 6 classes: murder, Class X (rape, deviate sexual assault, aggravated kidnapping for ransom, home invasion), Class 1 (aggravated kidnapping not for ransom, armed robbery), Class 2 (voluntary manslaughter, kidnapping, robbery, burglary, arson), Class 3 (involuntary manslaughter, aggravated battery, forgery, perjury and theft, which includes fraud, extortion and embezzlement), and Class 4 (reckless homicide, bigamy, pandering, theft of a firearm). The statutes which guide the trial judge in sentencing specify the following penalties. For murder, not less than 20 nor more than 40 years; if accompanied by aggravating circumstances, a term of natural life may be imposed; if unaccompanied by mitigating circumstances, the death penalty may be imposed. For a Class X felony, not less than 6 nor more than 30 years; for a Class 1 felony, not less than 4 nor more than 15 years; for a Class 2 felony, not less than 3 nor more than 7 years; for a Class 3 felony, not less than 2 nor more than 5 years; and for a Class 4 felony, not less than 1 nor more than 3 years. The trial judge is required to specify his reasons for imposing the particular sentence, including any mitigating

or aggravating factors which were considered. Although sentences may not be extended beyond those authorized by the statutes, the judge may elect to reduce the sentence to probation, conditional discharge, periodic imprisonment or a fine (Source: Smith-Hurd Illinois Annotated Statutes, Chapter 38, Section 1005-8-1).

There is no observable agency contract between society and juries; that is, society has no means of disciplining juries who disregard or subvert the law. To a great extent this is true of judges as well. According to the Smith-Hurd Illinois Annotated Statutes, "a judgment as to proper sentence depends upon many factors, such as defendant's prior record, credibility, demeanor, general moral character, mentality, social environment, habits and age" (Section 1005-8-1, Note 13, People v. Cozzi, 1981). However, the exact weights to be placed upon various matters relevant to sentencing are up to the trial judge--"It is not the function of appellate court to serve as sentencing court, and it will not substitute its judgment for that of trial court merely because it would have balanced appropriate factors differently" (Section 1005-5-3, Note 501, People v. Pace, 1981). All in all, the case for substantial discretion seems compelling.[3]

3. THE MODEL

The main point of this paper can be made in the context of a simple model in which an individual has a single opportunity to commit a crime of variable severity.[4] We assume a three-stage process in which society allocates resources to apprehension, individuals decide

upon the severity of their crime and subsequently society, through a representative such as a judge or jury, imposes sanctions in the form of prison sentences upon offenders who are apprehended and convicted. For our basic model, we assume that all members of society have the same preferences, so it is irrelevant whether society's representative is a judge or jury.[5] This structural detail will turn out to be important when we consider the effects of repeated play. Since society's representative moves last in this game, its preferred sentence length will generally depend upon the severity of the crime and other characteristics of the offender. Individuals, anticipating the behavior of society's representative, choose their preferred crime. Finally, anticipating the behavior of both criminals and its own representative, society allocates resources to apprehension. We model the resulting probability of apprehension as independent of the severity of the crime and the characteristics of the criminal. This corresponds to choosing the extent of police presence, with apprehension being an increasing function of the number of police. Alternatively, one could model the allocation of resources to apprehension taking place after a crime has occurred, but without knowledge of the offender's characteristic. Each of these models has some merit and deserves investigation. To make our point in the simplest possible context, we assume that society moves first, choosing a constant probability of apprehension, denoted by p.

Given the timing assumptions of our model, and the requirement of subgame perfection, it is appropriate to consider first the last

stage, in which a representative of society (i.e., a judge or jury) is assumed to choose its preferred sentence length, given that a crime has been committed and the offender has been apprehended and convicted. Social utility following apprehension is assumed to be a function of the length of the sentence, the severity of the crime and the costs of imprisonment.[6] Let x denote the length of a sentence and y the severity of a crime. We assume that individuals are indexed by a characteristic q which may best be interpreted as an indicator of the private and social cost of imprisonment. For example, q might be thought of as the criminal's alternative wage in lawful activity. Although society need not attribute the same disutiity to a criminal's incarceration as does the criminal himself, it is reasonable to include in the social utility function some measure of the loss of foregone lawful services of the criminal while in prison. We assume that this utility function is separable into a benefit function b(x,y) and a cost function c(x,q). The benefit function may represent utility derived from vengeance, retribution, or compensation of the specific victim or society in general (which is also to be considered a victim). The cost function summarizes opportunity costs and incarceration costs. Letting subscripts denote partial derivatives, we assume that $b(0,y) = b(x,0) = b_x(x,0) = 0$: that is, society derives no benefit (net of possible property recovery) from apprehension and conviction if no penalty is imposed; moreover, it derives no benefit from imprisoning individuals who have committed no crime. In addition, we assume that b(x,y) is twice continuously differentiable

with $b_{xx} < 0$, $b_y > 0$, and $b_{xy} > 0$ for $x,y > 0$: the benefit function is strictly concave in sentence length, and both the benefit and the marginal benefit b_x increase with the severity of the crime. Finally, we assume that for each $y > 0$, there exists $\hat{x}(y) \leq \infty$ such that $b_x(x,y) > 0$ for $x < \hat{x}(y)$, $b_x(x,y) = 0$ for $x = \hat{x}(y)$ and $b_x(x,y) < 0$ for $x > \hat{x}(y)$. The value $\hat{x}(y)$ maximizes the benefits from incarceration (without consideration of any associated costs) and is thus interpretable as the "just" or "optimal retributive" sentence.[7] Our assumptions regarding b_{xx} and b_{xy} imply that $\hat{x}'(y) > 0$; the optimal retributive penalty increases with the severity of the crime.

Of course, the sentence which is actually imposed may deviate from the optimal retributive sentence because society must rationally consider the costs of sentences as well as the benefits. We assume that $c(x,q)$ is twice continuously differentiable with $c_x > 0$, $c_{xx} \geq 0$, $c_q > 0$, and $c_{xq} > 0$. That is, costs increase at a nondecreasing rate with sentence length, and both total and marginal cost c_x increase with q. We assume that society bears no cost if it imposes no penalty; that is, $c(0,q) = 0$. Society's representative is assumed to choose the sentence length x so as to maximize social benefits net of costs, denoted $J(x;y,q)$.[8] We will refer to the solution to this problem as the credible sentence schedule, and will denote it by $x^*(y,q)$. Since $J(x;y,q)$ is strictly concave in x, if $x^*(y,q)$ is interior, then it is the unique solution of

$$J_x(x;y,q) = b_x(x,y) - c_x(x,q) = 0. \qquad (1)$$

Alternatively, if $J_x(x;y,q) < 0$ for all x given (y,q), then $x^*(y,q) = 0$. Given our assumptions on the benefit and cost functions, it is easy to show that $x^*(0,q) = 0$ for all q; that is, individuals who commit no crime receive no punishment. When $x^*(y,q)$ is positive so is b_x, implying that $x^*(y,q) < \hat{x}(y)$ for all (y,q); the credible sentence is always less than the optimal retributive sentence. For $x^*(y,q) > 0$,

$$\partial x^*(y,q)/\partial y = -b_{xy}/[b_{xx} - c_{xx}] > 0; \text{ and}$$
$$\partial x^*(y,q)/\partial q = -c_{xq}/[b_{xx} - c_{xx}] < 0.$$

That is, $x^*(y,q)$ varies directly with the severity of the crime y and inversely with the opportunity cost parameter q. Subgame perfection implies that the credible sentence schedule cannot depend upon the probability of apprehension; that is, $x^*(y,q)$ does not depend upon p. The classic result from the precommitment approach--that one can substitute more severe penalties for a lower probability of apprehension--is not characteristic of the model without precommitment.

The fact that x^* increases with y means that the credible sentence schedule is characterized by a "marginal penalty" of the sort called for by Stigler (1970) in a precommitment model. One only needs a marginal penalty if not all crime is deterred under the optimal enforcement policy. While this is not true in some classic instances of the precommitment model (Stern, 1978), it will be true in our model because not all crime can be credibly deterred. Given that some crime

will occur, a marginal penalty can at least reduce the extent to which more severe crimes are committed.

The inverse relationship between x^* and q has several possible interpretations; if q is thought of as a level of skill or alternative wage, then one could interpret this result as implying that a physician defrauding Medicaid would receive a shorter sentence than a welfare cheater of the same magnitude. Similarly, a middle- or upper-class murderer would receive a shorter sentence than his lower-class counterpart. If the alternative wage declines with the extent of prior criminal activity, then (all else equal) the criminal with the longer record will receive a longer sentence.

Now consider the problem of an individual who is contemplating criminal behavior. Recall that p is the probability of apprehension, which for simplicity is taken to be constant. What is the optimal crime y for an individual of type q? Let $U(y;q)$ denote the expected utility of an individual of type q who commits a crime y, and anticipates a sentence of $x^*(y,q)$ if apprehended. We assume that $U(y;q)$ takes the form

$$U(y;q) = pv(x^*(y,q),q) + (1 - p)u(y,q), \tag{2}$$

where $v(x^*(y,q),q)$ measures utility in the event of apprehension and $u(y,q)$ denotes utility if not apprehended. This specification seems appropriate in cases of property crime with recovery. Alternatively, for property crime without recovery or for crimes like assault, rape and murder, a more appropriate formulation would be

$$U(y;q) = pv(x^*(y,q),q) + u(y,q). \tag{3}$$

That is, the criminal enjoys the utility $u(y,q)$ simply by committing the crime and suffers the utility $v(x,q)$ only if apprehended. For concreteness, we will use the formulation in equation (2). It is straightforward to replicate the analysis using the formulation in equation (3).

We assume that $v(x,q)$ is twice continuously differentiable with $v_x(x,q) < 0$, $v_{xx}(x,q) < 0$ and $v_{xq}(x,q) < 0$: utility decreases at an increasing rate with the length of sentence, and the marginal utility of sentence x decreases with the individual characteristic q. The utility function $u(y,q)$ is assumed twice continuously differentiable with $u_y(y,q) > 0$, $u_{yy}(y,q) \leq 0$ and $u_{yq}(y,q) \leq 0$: the utility associated with a successful crime increases at a nonincreasing rate with y, and the marginal utility of crime y decreases with the characteristic q. This is consistent with the interpretation of q as an alternative wage. In order for this specification to make sense, we also require that $v(0,q) = u(0,q) = w(q)$; that is, the criminal suffers no disutility if he receives a sentence of length 0, and enjoys no additional utility if he commits no crime. His utility in the absence of criminal activity is measured by $w(q)$, where $w(q)$ is an increasing and concave function.

The individual of type q chooses y to maximize his or her expected utility, yielding the following first- and second-order necessary conditions for an interior maximum.

$$pv_x(x^*,q)\partial x^*/\partial y + (1 - p)u_y(y,q) \leq 0 \qquad (4)$$

$$pv_{xx}(\partial x^*/\partial y)^2 + pv_x(\partial^2 x^*/\partial y^2) + (1 - p)u_{yy} < 0. \qquad (5)$$

We will assume that equation (5) holds. If (4) holds with a strict inequality for all y, then the optimal severity is $y^*(q,p) = 0$; that is, no crime is committed. For example, this might be the case if p is sufficiently close to 1 and x^* is sufficiently sensitive to y in a neighborhood of $y = 0$. When the equilibrium crime is nontrivial, solving equation (4) with equality yields $y^*(q,p)$; then equations (4) and (5) imply that

$$\text{sgn } \partial y^*/\partial p = \text{sgn}\{v_x(\partial x^*/\partial y) - u_y\} < 0; \text{ and}$$

$$\text{sgn } \partial y^*/\partial q = \text{sgn}\{pv_{xq}(\partial x^*/\partial y) + pv_{xx}(\partial x^*/\partial q)(\partial x^*/\partial y) + pv_x(\partial^2 x^*/\partial q \partial y) + (1 - p)u_{qy}\}.$$

Thus the severity of crime depends inversely upon the probability of apprehension p. The dependence of severity upon the characteristic q is more complicated, involving four distinct expressions; the first and last are negative under our assumptions and previous results; the second is positive, and the third is ambiguous without further restrictions. Thus there may be a nonmonotonic relationship between the characteristic q and the severity of the associated crime. This possibility arises because of the anticipated dependence of x^* on q.

4. EXAMPLE

The following example is consistent with the assumptions of our basic model and shows that a solution to the two-stage problem can exist. More general existence questions will not be addressed here. For computational convenience, we assume that ethical considerations do not limit the length of sentence (that is, $\hat{x}(y) = \infty$ for all y) and we ignore discounting.

Suppose that the potential criminal works a day job and receives a wage of q; he contemplates how much to steal by night. Let y denote the magnitude of the theft. Once the criminal is apprehended and convicted, and assuming that y is recovered, the benefits function b(x,y) represents the extent to which imprisonment "compensates" society and the victim for the forced transfer. A plausible form for the benefit function is $b(x,y) = \alpha y x^{\beta}$, where $\beta \in (0,1)$. Let $c(x,q) = (c + q)x$; the parameter c represents incarceration costs per period and q represents society's opportunity cost of incarceration (lost wages from lawful activity). Maximizing the difference between these functions yields the credible penalty schedule $x^*(y,q) = [\alpha\beta y/(c + q)]^{1/(1-\beta)}$. Sentence length decreases with the cost parameters c and q and increases with the index of severity y. It also increases with the parameter α, which is a measure of the saliency of crime in general.

Consider now the criminal's decision problem. Assuming that the criminal has a linear utility function over his lifetime wealth, a sentence of x years yields a utility of $v(x,q) = q - qx$, while

successful theft of y yields utility of $u(y,q) = q + y$. If p is the probability of apprehension, and the penalty schedule $x^*(y,q)$ is anticipated, maximization of expected utility $U(y,q)$ implies that

$$y^*(q,p) = [(1 - p)(1 - \beta)/qp]^{(1-\beta)/\beta}[(c + q)/\alpha\beta]^{1/\beta}.$$

The optimal severity y^* decreases with the saliency of crime α, and with the probability of apprehension p; it increases with the cost parameter c. More interesting is the result that $\partial y^*/\partial q$ is positive, zero, or negative as q exceeds, equals or is less than $c(1 - \beta)/\beta$. Thus there is a nonlinear relationship between the alternative wage and the severity of crime. Those with relatively high and relatively low alternative wages commit the more serious crimes, with those in the middle choosing less serious crimes. Those with a low alternative wage are not deterred by long sentences at low cost; similarly, those with a high alternative wage are not deterred by the high cost because they expect and receive short sentences.

Although we have not been able to establish the following as a general property of equilibrium in our model, an interesting result emerges from consideration of this example. Despite the fact that sentence length is discriminatory, the total equilibrium penalty in terms of wealth or disutility is not. That is, if two types q_1 and q_2 commit the same crime y, then $q_1x^*(y,q_1) = q_2x^*(y,q_2)$. To see this, note that two types q_1 and q_2 commit the same crime if and only if $y^*(q_1,p) = y^*(q_2,p)$. Algebraic manipulation of this equality shows that it implies $q_1x^*(y^*,q_1) = q_2x^*(y^*,q_2)$. However, it is not true

that $q_1x^*(y,q_1) = q_2x^*(y,q_2)$ if q_1 and q_2 might (in equilibrium) choose different crimes. Thus all criminals who are observed to have committed the same crime will suffer equal punishment if apprehended. Whether this result extends to more general cases is not obvious, but risk neutrality on the part of the criminal is inessential; the same analysis applies to the constant absolute risk aversion utility function over wealth w given by $u(w) = -\exp[-w]$. What is clear is that if the total (equilibrium) penalty is to be nondiscriminatory, then the sentence length itself must discriminate among criminals with different values of q who commit the same crime y.

5. ALLOCATION OF RESOURCES TO APPREHENSION

We assume that the allocation of resources to apprehension generates a uniform probability of apprehension p, and is done before the commission of any crimes. Thus society is able to influence the severity of equilibrium crime because the severity of crime function $y^*(q,p)$ depends upon the parameter p. Similarly, the equilibrium sentence received by the individual of type q, $x^{**}(q,p) = x^*(y^*(q,p),q)$, will depend upon p through $y^*(q,p)$. Let F(q) denote the frequency distribution of the index q over its support $[\underline{q},\bar{q}] \subset (0,\infty)$.

Let D(y) denote the social disutility associated with the crime y; we assume that $D(\cdot)$ is twice continuously differentiable with $D' > 0$ and $D'' > 0$. There is some controversy over the inclusion of the criminal's utility from crime in the social utility function.

Becker (1968) discusses "the social value of the gain to offenders," and Polinsky and Shavell (1979) use a utilitarian social welfare function which results in a problem more reminiscent of external diseconomies than of criminal activity. Stigler (1970, p. 527) argues that there is little evidence "that society sets a positive value upon the utility derived from a murder, rape or arson." While Posner (1985, p. 1197) is willing to admit the criminal's utility from crime into the social utility function, he argues that it is almost surely outweighed by the disutility of the victim. "Now as a matter of fact it is a pretty safe empirical guess that most such conduct does create net disutility . . . it is unlikely that every disutile experienced by the wretched victim confers an equal and opposite utile on the offender." Following Posner, we assume that every crime y generates net disutility to society (where "net" means net of both the criminal's utility from crime and society's utility from punishing the apprehended offender). That is, we assume that $D(y) > R(y) + J(x^*(y,q);y,q)$ for all (y,q), where $R(y)$ summarizes the extent of recovery in the case of property crimes.

Anticipation of the equilibrium relationships $x^*(y,q)$ and $y^*(q,p)$ implies that society expects to suffer a loss of $D(y^*(q,p))$ for each criminal of type q, and to recoup a gain of $R(y^*,q,p) + J(x^{**}(q,p);y^*(q,p),q)$ with probability p. If $k(p)$ denotes the cost of sustaining the probability p of apprehension, then expected social utility can be written as follows.

$$ESU(p) = -\int D(y^*(q,p))dF(q) + p\int[R(y^*) + J(x^{**}(q,p);y^*(q,p),q)]dF(q) - k(p).$$

Assuming that x^*, y^*, and p^* are all interior, the optimal value p^* satisfies[9]

$$ESU'(p^*) = -\int D'(y^*)(\partial y^*/\partial p)dF(q) + \int[R(y^*) + J(x^{**},y^*,q)]dF(q) + p^*\int[R'(y^*) + J_y(x^{**},y^*,q)](\partial y^*/\partial p)dF(q) - k'(p^*) = 0. \quad (6)$$

Note that $dJ/dp = J_x(\partial x^{**}/\partial p) + J_y(\partial y^*/\partial p) + J_p$; the first term is zero whenever $x^* > 0$, and the last term is always zero. This accounts for the term $J_y(\partial y^*/\partial p)$ in the expression $ESU'(p^*)$.

This equation implies that there are two sources of benefits from increasing the probability of apprehension which must be balanced against two sources of cost. The first term represents the benefits from increased deterrence. A higher probability of apprehension induces criminals to choose less severe crimes, which reduces the social loss from crime. The second term summarizes the benefits derived from increased apprehension and punishment. The third term, which is negative (and is thus a cost of increased apprehension), describes the loss of utility associated with punishing lesser crimes. That is, because individuals are induced to commit less serious crimes, society derives lower utility from apprehending and punishing these criminals. The last term is the marginal resource cost of apprehension.

We have already discussed the possibility that police discretion may result in a reallocation of resources across crimes which would result in a nonconstant probability of apprehension.

Another source of variation in the likelihood of apprehension is that the same amount of effort devoted to various crimes can result in more or less apprehension. Typically the perpetrators of more serious crimes are more easily apprehended. More serious crimes tend to involve the victim directly, leaving more clues and often eye witnesses. This could be modeled by specifying a given level of enforcement effort e, which generates a probability p(y,e) of apprehension for crime y. Plausible assumptions regarding p(y,e) include $p(0,e) = p(y,0) = 0$, $p_y > 0$, $p_e > 0$, and $p_{ey} > 0$. That is, respectively, if no crime is committed, the criminal cannot be apprehended; if no effort is expended, no criminals can be apprehended; perpetrators of more serious crimes are easier to apprehend; an increase in effort increases the probability of apprehension and this increase is greater the more serious is the crime. An example is $p(y,e) = ey/(1 + y)$. In this case, the criminal chooses y to maximize his expected utility, anticipating the effect of his choice upon p(y,e) and $x^*(y,q)$. Again this yields a crime severity function $y^*(q,e)$ which may be nonmonotonic in q, and which now depends inversely upon enforcement effort e. Society is assumed to choose enforcement effort e to balance the benefits of greater deterrence and apprehension against the costs of enforcement.

6. DETERRENCE-ENHANCING STRATEGIES

Recall that since we have assumed that $D(y) > R(y) + J(x^*(y,q);y,q)$ for all (y,q), in our model society would

in principle prefer to deter all crime.[10] However, the discretion to determine conviction and/or sentencing after a crime has been committed will tend to restrict society's ability to do so. To see this, note that for each (y,q,p) triple, there is a shortest y-deterring sentence $x^o(y,q,p)$ which is defined implicitly by

$$pv(x^o,q) + (1 - p)u(y,q) = pv(0,q) + (1 - p)u(0,q) = w(q).$$

That is, $x^o(y,q,p)$ is the length of sentence which equates the expected utility which type q derives from crime y (given the probability of apprehension p) to the utility derived from no crime w(q). It is straightforward to verify that our previous assumptions imply that x^o varies directly with the severity of crime y and inversely with the probability of apprehension p.[11] For our example, $x^o(y,q,p) = (1 - p)y/pq$. Thus if the schedule x^o could be enforced, no crimes would be committed. Moreover, in this case society faces the usual incentives to lower p. which is costly, and to compensate by raising $x^o(y,q,p)$, which is costless since there will be no crimes to punish.

Despite the attractiveness of the implied outcome (a crime-free society), discretion at the point of sentencing renders it unattainable. Yet there are opportunities, both formal and informal, for society to improve its overall utility by enhancing the deterrence features of its law enforcement policies. For example, one opportunity involves a mandatory sentence schedule linking sentences to the severity of the crime and the characteristics of the offender.

The latter link is crucial; there may exist no mandatory sentencing policy which is both independent of q and improves society's overall welfare. Moreover, the choice of the correct mandatory sentence schedule may be quite complicated. For example, suppose that the deterring sentence schedule $x^{O}(y,q,p)$ were made mandatory, but a jury determines guilt or innocence; such a policy will be at least partially thwarted by the jury's refusal to convict.

This phenomenon, which is known as "jury nullification," is reasonably well-documented. For example, Kalven and Zeisel (1966, Ch. 21) offer some contemporary evidence as well as the following historical note (p. 311). "A memorable chapter in the history of the English jury concerns its response to excessive punishment in the early nineteenth century when England had an incredible list of some 230 capital offenses. The jury then felt the death penalty so disproportionate for most crimes that it conspicuously refused to convict. Finally, in 1819, the bankers themselves petitioned Parliament to remove the death penalty from the crime of forgery, since it had become almost impossible to obtain a conviction for that crime." Jackson (1984, pp. 152-53) describes a relatively recent mandatory sentencing experiment. "In the early 1960s the federal government introduced penalties of mandatory life sentences without possibility of parole or suspension for transferring narcotics: you didn't even have to sell the stuff; a junkie could give a hit to a sick pal and be culpable. Few dope dealers of any significance were punished under that legislation, though there were a number of

prosecutions. The first man convicted and sentenced to life was Gilbert Mora Zaragoza, a junkie with low grades on his intelligence test scores who was asked by a close friend to let him have some dope because the friend was going into withdrawal. . . . There were a few more such grossly irrelevant convictions and sentencings, and then the number of convictions under that legislation began falling off at an alarming rate. . . . Juries simply refused to believe that transferring a cap or two to a pal earned life in the penitentiary. Juries engaged in what the courts call <u>jury nullification</u>: even in light of uncontrovertible evidence, the juries brought in not-guilty verdicts anyway; they found the law unworthy of the defendants. Judges granted motions to suppress they previously would have rejected out of hand. The most immediate effect of the federal legislation was that federal prosecutors were so fearful of losing even very good cases in courts that they began accepting outrageously reduced charge bargains. Subsequently Congress repealed the penalties, went back to what seemed more reasonable sentences, and juries once again started delivering the usual guilty verdicts."

Nullification behavior need not be confined to juries. Wilson (1975, p. 187) remarks that "in Great Britain, where judges had less discretion in imposing the death penalty than they do in the United States, the number of murderers found insane—and so spared the gallows--dropped sharply after the death penalty was abolished in 1965. It is hard to believe that there were fewer insane persons in Britain after abolition of the death penalty; what apparently happened

was that the authorities no longer felt it as necessary to protect the accused from penalties when the penalty was no longer death. No one should assume that any judicial outcome can be made truly 'mandatory'-- discretion removed from one place in the criminal justice system tends to reappear elsewhere in it."

Within the context of our formal model, given the schedule $x^o(y,q,p)$, the judge or jury will convict q of y if and only if $J(x^o(y,q,p);q,y) \geq J(0;q,y)$; otherwise the criminal will be acquitted. Since $J_{xx} < 0$, there is a longest sentence $\tilde{x}(y,q)$ which will result in conviction; the sentence schedule $\tilde{x}$ will be referred to as the harshest convictable sentence schedule.[12] For our example (assuming y is recovered even if the offender is not convicted), $\tilde{x}(y,q) = [\alpha y/(c + q)]^{1/(1-\beta)}$. The judge or jury will refuse to convict an individual of type q of a crime of severity y whenever $x^o(y,q,p) \geq \tilde{x}(y,q)$; that is, whenever $(1 - p)y/pq \geq [\alpha y/(c + q)]^{1/(1-\beta)}$. This inequality holds whenever the probability of apprehension p, the saliency of crime α, the opportunity cost of incarceration q or the severity of the crime y is sufficiently small, or when either of the cost parameters c or q is sufficiently large. If the sentence schedule x^o were to be imposed, what would actually be observed is x^o when $x^o \leq \tilde{x}$ and 0 otherwise. There is a unique value of y for each q which equates the two schedules; call this $y^o(q,p)$. Then examination of the criminal's expected utility function for the example implies that the crime function will be $y^o(q,p)$. It is straightforward to show that for $p < q$, $0 < y^o(q,p) < y^*(q,p)$ for all q. Thus the naive

deterrence strategy which consists of legislating the schedule x^o does work to some extent, but it is limited by the discretionary power to refuse to convict. Moreover, this policy need not improve overall social utility because although deterrence is enhanced, no criminals are ever penalized and this reduces social utility.[13] The schedules x^o, x^*, and $\tilde{x}$ are illustrated in Figure 1.

Our basic model treats criminal activity and law enforcement as a "one-shot" game. Since sanctions are required to be credible, they cannot have a "pure deterrence" component; nevertheless, credible sanctions do deter to the extent that they induce individuals to commit less serious crimes. In reality, the law enforcement game is repeated with the same and/or different criminals, which suggests the possibility that current penalties might foster future deterrence. The assumption that each criminal has only one opportunity for crime does not rule out repeated games arguments, since these are equally applicable to a sequence of identical criminals, each of whom has only one opportunity for crime. However, if society's preferences are known, simple repetition of the one-shot game has as its only equilibrium an infinite sequence of one-shot equilibria; a more sophisticated institution is needed to sustain any other policy.

For example, repeated games arguments which rely on the use of "trigger strategies" can often render credible a policy which would not be credible in the absence of repetition (Friedman, 1971). Similarly, it is possible that incomplete information on the part of criminals may enable society to establish a "reputation for

deterrence" by repeatedly imposing the deterring sentence schedule (Kreps and Wilson, 1982b). However, these sorts of arguments are very sensitive to both the identity of society's representative and, in the latter case, to the assumed form of criminals' uncertainty.

For instance, suppose that society's representative is a jury which is responsible for both conviction and the extent of sanctions. Since a jury is convened once for the specific case, no single jury is engaged in the repeated game. But consider the following "trigger strategy": impose the schedule $x^m(y,q,p) = \max[x^o(y,q,p), x^*(y,q)]$ if it was imposed in the last case; otherwise impose the credible schedule $x^*(y,q)$. The schedule x^m completely deters crime, but allows society to choose its most preferred sentence when this exceeds the minimum deterring penalty x^o. Notice that once a jury fails to impose the deterring sentence x^m, society must switch to p^* and criminals to $y^*(q,p^*)$ in anticipation of the reversion to the credible sentence schedule x^*. If all juries follow this rule, the deterring policy (p,x^m) can be sustained in repeated play if and only if, for every conceivable case (y,q), the value of maintaining the trigger exceeds the value of reversion. That is,

$$J(x^m;y,q) + \left(\frac{\delta}{1-\delta}\right)(-k(p)) \geq J(x^*;y,q) + \left(\frac{\delta}{1-\delta}\right)ESU(p^*),$$

where δ is the jury's discount factor. Alternatively put, for each case (y,q), the cost of maintaining the trigger must be less than the associated benefit.

$$J(x^*;y,q) - J(x^m;y,q) \leq \left(\frac{\delta}{1-\delta}\right)[-k(p) - ESU(p^*)]. \qquad (7)$$

In evaluating the likelihood that x^m will be sustained, it is necessary to consider both its feasibility and its optimality. It seems unlikely that prospective jurors are sufficiently well-informed about past cases to use such a strategy; at the least they would have to reconstruct the previous case in order to determine whether or not the proper penalty was imposed (recall that they would need to know q as well as y and x). If such a policy were feasible, it is necessary to determine whether inequality (7) is satisfied. The right-hand-side of (7) is the net gain to society from deterring future crime, and it is bounded above as a function of p since p does not affect $ESU(p^*)$. The left-hand-side represents the current cost of deterring future crime; it is the difference between (ex post) utility under the credible and deterring sentence schedules. But the minimum deterring sentence x^o is that value of x such that $v(x^o,q) = [w(q) - u(y,q)(1-p)]/p$; moreover, $w(q) < u(y,q)$ for all $y > 0$. Thus the deterring schedule x^m requires unbounded disutility as p nears 0, even for small values of y. Recalling that the cost to society $c(x,q)$ reflects in some measure the cost $v(x,q)$ to the criminal, it is plausible that unbounded disutility to the criminal implies unbounded disutility for society. For example, suppose that on account of its very low probability of apprehension, the deterring sentence for bicycle theft is prolonged torture; even if a single imposition of this sentence would deter all future crime, society may reckon this too great a cost. Since the left-hand-side of (7) seems likely to be unbounded as p nears 0, the ideal precommitment policy of

a very small probability of detection and a very long sentence is unlikely to be sustainable by repeated play. Of course, the greater is p, the lower is the deterring schedule and the more likely it becomes that inequality (7) will be satisfied.[14]

One way for society to develop a reputation for deterrence in this context would involve assuming that all juries prefer (ex post) the deterring sentence $x^m(y,q,p)$ or all juries prefer the credible sentence $x^*(y,q)$, with the criminal being uncertain as to which is true. Then all juries are effectively alike, and the behavior of one jury can plausibly be taken to reflect on the common preferences of other juries. In this case, one is in a most favorable situation if a single round of play is sufficient to establish one's reputation. In this pleasant circumstance, the same inequality (7) must hold. Thus the reputation building model is subject to the same difficulties regarding optimality, and reputation building is even less attractive if more periods are needed to establish a reputation. A more plausible specification of criminal uncertainty would be that each jury prefers either the deterring sentence schedule $x^m(y,q,p)$ or the credible sentence schedule $x^*(y,q)$, with criminals being uncertain about the proportion which prefers the deterring schedule. In this case, the behavior of one jury cannot be taken as providing information about the preferences of juries as a whole, and no reputation for deterrence can be established.

If society's agent is a judge, the feasibility of using trigger strategies seems somewhat more likely. Judges might be

presumed to be better informed about the previous decisions of their colleagues than are juries. This information is necessary since no judge can punish himself; judges must be "paired" so each can punish his partner for failures to impose x^m. The optimality of sustaining x^m by means of trigger strategies is again determined by whether inequality (7) holds. Reputation building on the part of judges would rely on the assumption that some judges prefer (ex post) the deterring schedule while others prefer the credible schedule, with the criminal being uncertain about the relative magnitudes of these groups. This uncertainty might provide all judges with sufficient incentive to behave as if they preferred the deterring schedule so as to build a reputation for deterrence. Again the optimality of reputation building rests on the fact that the benefits of reputation formation must outweigh the costs. In the happy event that only one round is needed to establish a reputation, we again need only check that (7) holds. If more than one round is needed, (7) may hold and yet it need not be optimal to establish a reputation for deterrence. The essential element for sustaining deterrence through repeated play is the dependence of strategies upon the sentencing history; yet one of the most striking features of the existing criminal justice system is an almost complete lack of systematic sentencing review.

Finally, if individual members of society have diverse preferences regarding sentencing, one plausible means of precommitment is to appoint or elect judges with particular preferences, and simply allow them to exercise their discretion. This argument does not apply

quite as well to the selection of juries, since opposing counsel will each try to screen out jurors who favor the other side. Thus even if it is possible to bias systematically the selection of judges, the problem of jury nullification is likely to remain.

7. CONCLUSIONS

The precommitment approach to law enforcement is inappropriate as a positive theory of crime and punishment because it is inconsistent with the institutional structure of U.S. law enforcement. Moreover, under the assumption that all crime is potentially deterrable, it does not seem to generate accurate predictions regarding the extent of criminal activity and sanctions. Recently both Posner (1985) and Shavell (1985) have argued that crime exists because some individuals simply cannot be deterred with the number of undeterrables rising as the probability of apprehension falls. In this case, optimal sentences must reflect a balancing of the benefits of deterrence and the costs of actually imposing sanctions (including the possible wrongful punishment of innocents). This balance implies that sentence length should vary directly with the severity of crime and inversely with the costs of imprisonment and the probability of apprehension. Our alternative approach requires the penalties imposed on convicted offenders to be optimal given that the crime has been committed. This implies a sentence schedule which varies directly with severity and inversely with the costs of imprisonment; however, the credibility requirement severs the tie between the probability of

detection and the penalty which is characteristic of the precommitment approach. Moreover, even the less extreme sentences implied by Posner and Shavell's version of the precommitment approach will suffer a credibility problem given discretion at the point of sentencing.

FOOTNOTES

* We would like to thank Michael Graetz, Alvin Klevorick, and Sam Peltzman for helpful comments and discussions. Reinganum gratefully acknowledges the financial support of the Alfred P. Sloan Foundation, the Center for the Study of the Economy and the State and the University of Chicago's Graduate School of Business. Both authors acknowledge the support of National Science Foundation grant no. SES-8315422.

1. There are of course some exceptions; in particular, mandatory sentencing laws do exist. The effects of these laws will be discussed in more detail in Section 6.

2. See Polinsky and Shavell (1979) for a model in which some crime is socially desirable.

3. One might argue that judges are in fact faced with optimal agency contracts which, despite the appearance of discretion, induce them to choose in all cases the appropriate penalty. However, such a contract would require that judges be rewarded on the basis of the appropriateness of their decisions. But, according to one observer of the contemporary criminal justice system (Jackson, 1984, p. 169), "Judges are given enormous discretionary power by most state and federal laws so punishments will be appropriate to the crime and the defendant. But no measure of that appropriateness has ever been made, nor shall one be. That

is because, under the present system, the only definition of appropriateness has been the judge saying 'That's what I think he deserves.'"

4. This is in contrast to the approach in Becker (1968), in which the choice variable of the potential criminal is the frequency of a given crime.

5. There are many reasons why judges' and juries' preferences need not be consistent with social preferences (however determined). To isolate the effects of credibility, we will assume that all agents of society have identical preferences.

6. Given that the individual has only one opportunity for crime and has been apprehended and convicted, the requirement of subgame perfection implies that these are the only things on which social utility can depend. What happens at this point in the decision tree cannot depend upon anything that might or would happen at some other point.

7. Several authors have rejected the economic approach to explaining sanctions. Stern (1978) argues that the economic approach cannot explain why punishments are not raised (since they deter crime in the standard approach); instead he argues that society has a notion of the "just" punishment, which depends on the extent of the damage caused by the criminal. Wittman (1974) takes the optimal retributive punishment as given, and argues that justice

decreases as actual punishments deviate from the retributive optimum. Our analysis, which takes an explicitly economic approach, is designed to be consistent with the notion of a just or optimally retributive punishment. We provide an explanation of why punishments are not raised (unlimited punishments are not credible, and hence do not deter), and we argue that the credible sentence will be less than the retributive optimum because of the cost associated with sanctions.

8. We could have begun with the function $J(x;y,q)$, but separating the objective function into benefit and cost components is intuitive and without significant loss of generality.

9. To show that p^* can be interior, we compute an example based upon the analysis in Section 4 assuming $D(y) = y + \delta y^{\gamma}$ and $R(y) = y$. From Section 4, $J(x^*;y,q) = \eta(q)y^{\theta}$, where $\eta(q) = \alpha(1 - \beta)[\alpha\beta/(c + q)]^{\beta/(1-\beta)}$ and $\theta = 1/(1 - \beta)$. Assuming $\delta > \eta(0)$ and $\gamma > \theta$ implies that $D(y) > R(y) + J(x^*;y,q)$ and $D'(y) > R'(y) + J_y(x^*;y,q)$ for all (y,q). From Section 4, $y^*(q,p) \to \infty$ and $dy^*/dp \to -\infty$ for all q as $p \to 0$. Moreover, $R(y^*) + J(x^{**};y^*,q) \to \infty$ and $-D'(y^*) + p[R'(y^*) + J_y(x^{**};y^*,q)] \to -\infty$ as $p \to 0$. Thus $ESU'(0) > 0$ so long as $k'(0) < \infty$, implying that p^* is strictly positive for the example.

10. In this paper we take the usual economic approach in which preferences are defined only over the possible outcomes and not over the means by which the outcomes are achieved; that is, we do not discuss here the ethics of deterrence--punishing an offender severely in light of the low probability of apprehension in order to raise the expected penalty perceived by potential offenders--only its efficacy.

11. The expression $\partial x^o/\partial y$ has the same sign as $u_y > 0$. The expression $\partial x^o/\partial p$ has the same sign as $v(x^o,q) - u(y,q) < 0$ for $x^o, y > 0$. This follows from the relationships $v(x,q) < v(0,q) = u(0,q) < u(y,q)$ for $x,y > 0$. Finally, the expression $\partial x^o/\partial q$ has the same sign as $p(v_{qq} - u_q) + u_q - w'(q)$. Under the plausible assumption that $v_q < u_q \leq w'(q)$, the shortest y-deterring sentence x^o varies inversely with the characteristic q. Thus the key difference between x^o and x^* is that x^o varies inversely with p while x^* is independent of p.

12. It turns out that in our example, every type of criminal is indifferent between facing the credible schedule $x^*(y,q)$ and the harshest convictable schedule $\tilde{x}(y,q)$, although they choose different optimal crimes under these different schedules.

13. The best mandatory sentencing schedule is the one which maximizes society's overall welfare, subject to being convictable. That is, society chooses p and $x(y,q)$ to maximize an objective akin to the one in equation (6), subject to the constraints that the potential criminal chooses his preferred severity and that $x(y,q) \leq \tilde{x}(y,q)$ for all (y,q). This is consistent with the idea

that a legislative body chooses the penalty schedule, anticipating the possibility of jury nullification. Although it is relatively straightforward to state this problem, its solution is an open issue.

14. We know from the Folk Theorem that if it were feasible for juries to use trigger strategies, then the highest sustainable social payoff will typically exceed the payoff associated with the credible penalty schedule. How much better society can do in this case, and exactly what penalty schedule it would use are questions for future research.

REFERENCES

Becker, Gary S. "Crime and Punishment: An Economic Approach." Journal of Political Economy 76 (1968):169-217.

Block, M. K., and J. M. Heineke. "A Labor Theoretic Analysis of Criminal Choice." American Economic Review 65 (1975):314-325.

Brown, W. W., and M. O. Reynolds. "Crime and 'Punishment': Risk Implications." Journal of Economic Theory 6 (1973):508-514.

Friedman, James W. "A Noncooperative Equilibrium for Supergames." Review of Economic Studies 38 (1971):1-12.

Graetz, Michael J., Jennifer R. Reinganum, and Louis L. Wilde. "The Tax Compliance Game: Toward an Interactive Theory of Law Enforcement." Journal of Law, Economics and Organization 2 (1986):1-32.

Heineke, J. M. "A Note on Modeling the Criminal Choice Problem." Journal of Economic Theory 10 (1975):113-116.

Jackson, Bruce. Law and Disorder: Criminal Justice in America. Chicago: University of Illinois Press, 1984.

Kalven, Harry, Jr., and Hans Zeisel. The American Jury. Boston: Little, Brown and Company, 1966.

Kreps, David M., and Robert Wilson. "Sequential Equilibrium." Econometrica 50 (1982a):863-894.

_____. "Reputation and Imperfect Information." Journal of Economic Theory 27 (1982b):253-279.

Kydland, Finn, and Edward Prescott. "Rules Rather than Discretion: The Inconsistency of Optimal Plans." Journal of Political Economy 85 (1977):473-492.

Polinsky, A. Mitchell, and Steven Shavell. "The Optimal Tradeoff between the Probability and Magnitude of Fines." American Economic Review (1979):880-891.

Posner, Richard A. "An Economic Theory of the Criminal Law." Columbia Law Review 85 (1985):1193-1231.

The Roscoe Pound--American Trial Lawyers Foundation. The American Jury System. Annual Chief Justice Earl Warren Conference on Advocacy in the United States, 1977.

Pyle, David J. The Economics of Crime and Law Enforcement. London: Macmillan Press, 1983.

Selten, Reinhard. "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games." International Journal of Game Theory 4 (1975):25-55.

Shavell, Steven. "Criminal Law and the Optimal Use of Nonmonetary Sanctions as a Deterrent." Columbia Law Review 85 (1985):1232-1262.

Smith-Hurd Illinois Annotated Statutes. Criminal Law and Procedure, Chapter 38, 1982.

Stern, Nicholas. "On the Economic Theory of Policy Towards Crime." In Economic Models of Criminal Behavior. Edited by J. M. Heineke, 123-152. New York: North-Holland, 1978.

Stigler, George J. "The Optimum Enforcement of Laws." Journal of Political Economy 78 (1970):526-536.

Wilson, James Q. Thinking about Crime. New York: Basic Books, 1975.

Wittman, Donald. "Punishment as Retribution." Theory and Decision 4 (1974):209-237.

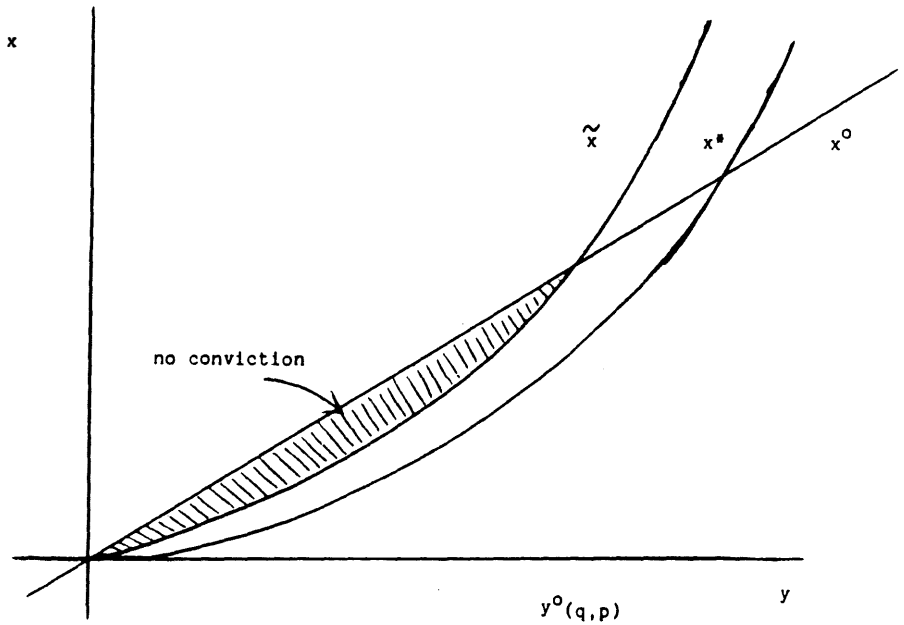

Figure 1